SUSMAN GODFREY L.L.P.

November 21, 2024

Hon. Ona T. Wang
Southern District of New York
500 Pearl Street
New York, New York 10007

> Re: *The New York Times Company v. Microsoft Corporation, et al.*,
> No.: 23-cv-11195-SHS: Opposition to OpenAI's Motion to Compel (Dkt. 310)

Dear Magistrate Judge Wang:

Plaintiff The New York Times Company ("The Times") opposes OpenAI's November 18 motion to compel the production of (1) documents and communications regarding The Times's "alleged market harm," and (2) documents and communications regarding "the markets for licensing the Times's works." Dkt. 310 at 1. OpenAI's motion is duplicative of a motion filed prior to the previous discovery conference, and is in any case unnecessary, as The Times has already agreed to produce numerous documents related to revenue, traffic, market harm, and licensing that are more than sufficient to address OpenAI's legitimate needs for discovery.

1. **OpenAI's November 18 Motion Is Duplicative of Its Pending October 22 Motion (Dkt. 276).**

Prior to the Court's October 30 discovery conference, OpenAI moved to compel the production of additional documents responsive to its RFPs 35 and 37, which seek documents concerning "analyses in readership and subscription trends" and analyses of the "performance of content published on The Times's website," respectively. Dkt. 276 ("the October 22 motion"). As discussed at the October 30 discovery conference, The Times has agreed to produce numerous categories of documents related to revenue, traffic, readership, and subscriptions, including analyses and reports regarding whether decreases in traffic to Times websites were attributable to Defendants' generative AI products or other market forces. *See* Dkt. 293 at 1 (summarizing related documents that The Times agreed to produce). Discussing this first motion, the Court expressed skepticism that The Times's offers were insufficient. *See* Hearing Tr. at 82:12–14 ("I'm not sure that asking for separate analyses on these other issues is necessary if the analyses of the AI influence have already taken that into account."). The present motion should be denied for the reasons set forth in The Times's prior briefing—namely, The Times has agreed to produce the relevant documents, and OpenAI's request is enormously burdensome because it seeks granular data such as real-time monitoring of individual articles on The Times's website that is simply untethered from the issues in this case. *See* Dkt. 293.

2. **The Times Has Already Agreed to Produce All Relevant Documents Concerning Decreases in Revenue and Web Traffic.**

OpenAI seeks documents concerning The Times's "alleged market harm, i.e., decreases in revenue and traffic, and analyses of the causes." Dkt. 310 at 1. As outlined in our response to the October 22 motion, The Times has already agreed to produce a significant volume of documents related to The Times's losses, including documents and data related to the magnitude of losses,

their causes, and the markets in which they occurred. *See* Dkt. 293 at 1. Since the October 30 hearing, in response to additional RFPs from both OpenAI and Microsoft, The Times has agreed to even more related discovery, committing to produce additional categories of documents related to 1) web traffic and/or revenue for Wirecutter, The Athletic, and NYT Cooking, 2) subscriptions, page views, user visits, and revenue from 2018 to 2024, and 3) analyses and reports regarding whether decreases in traffic to Times websites are attributable to GPT services versus other market forces and factors.[1]

All of these documents, and in particular this third category of documents concerning the causal relationship between Defendants' products and decreases in traffic to Times sites, will be responsive to OpenAI's fair use defense, sufficiently allowing "OpenAI to probe the Times's . . . claim that OpenAI's tools were the cause" of any decreases in traffic and revenue. Dkt. 310 at 2. OpenAI's requests for all documents and communications concerning any "other . . . causes of the alleged declines," *id.*, in traffic and revenue are unduly burdensome, including because these requests would capture a vast number of trivial daily communications between Times employees relating to *any* decrease in revenue or web traffic, regardless of the scale, cause, or relation to Defendants' products or the issues in this litigation.

3. **The Times Has Already Agreed to Produce All Relevant Documents Concerning the Market for Times Works.**

OpenAI also seeks documents concerning the "markets for licensing The Times's works." Dkt. 310 at 1. Here again, The Times has already agreed to produce numerous categories of relevant documents, both in response to the RFPs on which OpenAI seeks to compel production— RFPs 18, 33, and 160–61—and in response to other related requests regarding the market for Times Works. In response to RFPs 18, 33, and 160–61, The Times has agreed to produce:

> (1)   licensing agreements related to the Asserted Works (regardless of whether those licenses relate to generative AI);
>
> (2)   documents sufficient to show all attempts to license the asserted works for training or use in generative AI; and
>
> (3)   all communications related to executed and contemplated agreements to license Times works to train generative AI models.

And in response to other requests concerning the impact of Defendants' products on the market for the Asserted Works, The Times has further agreed to produce:

> (1)   all documents and communications regarding attempts to license or negotiate Times works to Defendants for use in their generative AI tools (OAI RFP 40);
>
> (2)   documents sufficient to show the "well-established market for The Times to provide paid access to and use of the Asserted Works" (OAI RFP 134 and Microsoft RFP 57);
>
> (3)   documents and communications regarding any lost market value of the asserted works as a result of the harm alleged in the complaint (OAI RFP 135);

---

[1] A chart summarizing precisely what The Times has agreed to produce in each category is attached as Exhibit A.

(4) analyses and reports regarding whether decreases in online traffic to Times websites are attributable to Defendants' products or other market forces (OAI RFP 145);

(5) documents and communications regarding any injury or harm suffered as a result of the conduct alleged in The Times's (OAI RFP 26); and

(6) documents concerning past and projected compensation from licensing the asserted works (Microsoft RFPs 6, 7).

But OpenAI seeks more, demanding all internal communications regarding licensing the Asserted Works for purposes other than generative AI model training, and (2) the existence (or lack thereof) of a "functioning market" for LLM-related licenses that are not related to generative AI training. Both requests are unduly burdensome, and OpenAI has provided no justification for why the documents that The Times has agreed to produce are inadequate to establish the market for licensing the Asserted Works and the effect of Defendants' products on that market.

OpenAI's request for internal communications regarding all licensing agreements is overly broad and unduly burdensome. The fair use analysis requires the Court to look at "the effect of [Defendants'] use upon the potential market for or value of the copyrighted work." 17 U.S.C. § 107(4). In evaluating the fourth fair use factor, courts examine "whether, if the challenged use becomes widespread, it will adversely affect the potential market for the copyrighted work." *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 613 (2d Cir. 2006). It is unclear how all of The Times's internal communications or negotiations about licensing agreements unrelated to any of the allegations in this case could be "highly probative under the fair use inquiry," Dkt. 310 at 3, particularly when The Times has already agreed to produce the licenses themselves.

And with respect to OpenAI's request for communications regarding the existence of a market for LLM-related licenses, The Times has already agreed to produce licensing agreements for the Asserted Works, documents sufficient to show the "well-established market for The Times to provide paid access to and use of the Asserted Works," and all documents and communications related to both executed and contemplated agreements to license Times works to train generative AI models. Together, this discovery will show the existence (or non-existence) of a market for LLM-related licenses where they are related to generative AI. Additional internal communications are unnecessary and unduly burdensome.

Finally, OpenAI's claim that "The Times has insisted that OpenAI produce communications regarding licenses and attempted licenses" misconstrues The Times's discovery requests to OpenAI. The Times has only requested that OpenAI produce documents and communications concerning licenses and data access agreements regarding issues related to this case, namely, the use of text data for generative AI model training and retrieval-augmented generation. OpenAI's overly broad request that the Times produce all licensing communications unrelated to the issues in this case is hardly "reciprocal." *See* Dkt. 310 at 3.

Given the duplicative nature of OpenAI's motion as well as the numerous categories of relevant discovery that The Times has already agreed to produce, The Times respectfully requests that the Court deny OpenAI's motion.

Respectfully submitted,

/s/ Ian B. Crosby

Ian B. Crosby

Susman Godfrey L.L.P.

/s/ Steven Lieberman
Steven Lieberman
Rothwell, Figg, Ernst & Manbeck

cc: All Counsel of Record (via ECF)